

#10



ENTERED

OIPE

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/819,104A

TIME: 15:03:47

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04122002\I819104A.raw

```

3 <110> APPLICANT: Chen, J. Don
5 <120> TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
6   AND USES THEREFOR
8 <130> FILE REFERENCE: UMG-030
10 <140> CURRENT APPLICATION NUMBER: 09/819,104A
11 <141> CURRENT FILING DATE: 2001-03-27
13 <150> PRIOR APPLICATION NUMBER: 60/193,138
14 <151> PRIOR FILING DATE: 2000-03-29
16 <160> NUMBER OF SEQ ID NOS: 6
18 <170> SOFTWARE: PatentIn Ver. 2.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 8686
22 <212> TYPE: DNA
23 <213> ORGANISM: Homo sapiens
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (157)..(7677)
29 <400> SEQUENCE: 1
30 gagtctttga ggacacagcc tcgctggagg cagtttctgg tgccagtgac ggggtggccc 60
32 gtgagctgat gacgaggact ggcttttaat ccttggtggt gattaagaga aagcttattg 120
34 gggcctggga gcagctcccc gcgaccccc accacc atg tcg ggc tcc aca cag 174
35                               Met Ser Gly Ser Thr Gln
36                               1           5
38 cct gtg gca cag acg tgg agg gcc act gag ccc cgc tac ccg ccc cac 222
39 Pro Val Ala Gln Thr Trp Arg Ala Thr Glu Pro Arg Tyr Pro Pro His
40           10           15           20
42 agc ctt tcc tac cca gtg cag atc gcc cgg acg cac acg gac gtc ggg 270
43 Ser Leu Ser Tyr Pro Val Gln Ile Ala Arg Thr His Thr Asp Val Gly
44           25           30           35
46 ctc ctg gag tac cag cac cac tcc cgc gac tat gcc tcc cac ctg tcg 318
47 Leu Leu Glu Tyr Gln His His Ser Arg Asp Tyr Ala Ser His Leu Ser
48           40           45           50
50 ccc ggc tcc atc atc cag ccc cag cgg cgg agg ccc tcc ctg ctg tct 366
51 Pro Gly Ser Ile Ile Gln Pro Gln Arg Arg Arg Pro Ser Leu Leu Ser
52 55           60           65           70
54 gag ttc cag ccc ggg aat gaa cgg tcc cag gag ctc cac ctg cgg cca 414
55 Glu Phe Gln Pro Gly Asn Glu Arg Ser Gln Glu Leu His Leu Arg Pro
56           75           80           85
58 gag tcc cac tca tac ctg ccc gag ctg ggg aag tca gag atg gag ttc 462
59 Glu Ser His Ser Tyr Leu Pro Glu Leu Gly Lys Ser Glu Met Glu Phe
60           90           95           100
62 att gaa agc aag cgc cct cgg cta gag ctg ctg cct gac ccc ctg ctg 510
63 Ile Glu Ser Lys Arg Pro Arg Leu Glu Leu Leu Pro Asp Pro Leu Leu

```

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/819,104A

TIME: 15:03:47

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04122002\I819104A.raw

64	105	110	115	
66	cga ccg tca ccc ctg ctg gcc acg ggc cag cct gcg gga tct gaa gac	558		
67	Arg Pro Ser Pro Leu Leu Ala Thr Gly Gln Pro Ala Gly Ser Glu Asp			
68	120	125	130	
70	ctc acc aag gac cgt agc ctg acg ggc aag ctg gaa ccg gtg tct ccc	606		
71	Leu Thr Lys Asp Arg Ser Leu Thr Gly Lys Leu Glu Pro Val Ser Pro			
72	135	140	145	150
74	ccc agc ccc ccg cac act gac cct gag ctg gag ctg gtg ccg cca cgg	654		
75	Pro Ser Pro Pro His Thr Asp Pro Glu Leu Glu Leu Val Pro Pro Arg			
76	155	160	165	
78	ctg tcc aag gag gag ctg atc cag aac atg gac cgc gtg gac cga gag	702		
79	Leu Ser Lys Glu Glu Leu Ile Gln Asn Met Asp Arg Val Asp Arg Glu			
80	170	175	180	
82	atc acc atg gta gag cag cag atc tct aag ctg aag aag aag cag caa	750		
83	Ile Thr Met Val Glu Gln Gln Ile Ser Lys Leu Lys Lys Lys Gln Gln			
84	185	190	195	
86	cag ctg gag gag gag gct gcc aag ccg ccc gag cct gag aag ccc gtg	798		
87	Gln Leu Glu Glu Glu Ala Ala Lys Pro Pro Glu Pro Glu Lys Pro Val			
88	200	205	210	
90	tca ccg ccg ccc atc gag tcg aag cac cgc agc ctg gtg cag atc atc	846		
91	Ser Pro Pro Pro Ile Glu Ser Lys His Arg Ser Leu Val Gln Ile Ile			
92	215	220	225	230
94	tac gac gag aac cgg aag aag gct gaa gct gca cat cgg att ctg gaa	894		
95	Tyr Asp Glu Asn Arg Lys Lys Ala Glu Ala Ala His Arg Ile Leu Glu			
96	235	240	245	
98	ggc ctg ggg ccc cag gtg gag ctg ccg ctg tac aac cag ccc tcc gac	942		
99	Gly Leu Gly Pro Gln Val Glu Leu Pro Leu Tyr Asn Gln Pro Ser Asp			
100	250	255	260	
102	acc cgg cag tat cat gag aac atc aaa ata aac cag gcg atg cgg aag	990		
103	Thr Arg Gln Tyr His Glu Asn Ile Lys Ile Asn Gln Ala Met Arg Lys			
104	265	270	275	
106	aag cta atc ttg tac ttc aag agg agg aat cac gct cgg aaa caa tgg	1038		
107	Lys Leu Ile Leu Tyr Phe Lys Arg Arg Asn His Ala Arg Lys Gln Trp			
108	280	285	290	
110	gag cag aag ttc tgc cag cgc tat gac cag ctg atg gag gcc tgg gag	1086		
111	Glu Gln Lys Phe Cys Gln Arg Tyr Asp Gln Leu Met Glu Ala Trp Glu			
112	295	300	305	310
114	aag aag gtg gag cgc atc gag aac aac ccc cgg cgg gcc aag gag	1134		
115	Lys Lys Val Glu Arg Ile Glu Asn Asn Pro Arg Arg Arg Ala Lys Glu			
116	315	320	325	
118	agc aag gtt cgc gag tac tac gag aag cag ttc cct gag atc cgc aag	1182		
119	Ser Lys Val Arg Glu Tyr Tyr Glu Lys Gln Phe Pro Glu Ile Arg Lys			
120	330	335	340	
122	cag cgc gag ctg cag gag cgc atg cag agg gtg ggc cag cgg ggc agt	1230		
123	Gln Arg Glu Leu Gln Glu Arg Met Gln Arg Val Gly Gln Arg Gly Ser			
124	345	350	355	
126	ggg ctg tcc atg tcg ccc gcc cgc agc gag cac gag gtg tca gag atc	1278		
127	Gly Leu Ser Met Ser Pro Ala Arg Ser Glu His Glu Val Ser Glu Ile			
128	360	365	370	

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/819,104A

TIME: 15:03:47

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04122002\I819104A.raw

```

130 atc gat ggc ctc tca gag cag gag aac ctg gag aag cag atg cgc cag 1326
131 Ile Asp Gly Leu Ser Glu Gln Glu Asn Leu Glu Lys Gln Met Arg Gln
132 375 380 385 390
134 ctg gcc gtg atc ccg ccc atg ctg tac gac gct gac cag cag cgc atc 1374
135 Leu Ala Val Ile Pro Met Leu Tyr Asp Ala Asp Gln Gln Arg Ile
136 395 400 405
138 aag ttc atc aac atg aac ggg ctt atg gcc gac ccc atg aag gtg tac 1422
139 Lys Phe Ile Asn Met Asn Gly Leu Met Ala Asp Pro Met Lys Val Tyr
140 410 415 420
142 aaa gac cgc cag gtc atg aac atg tgg agt gag cag gag aag gag acc 1470
143 Lys Asp Arg Gln Val Met Asn Met Trp Ser Glu Gln Glu Lys Glu Thr
144 425 430 435
146 ttc cgg gag aag ttc atg cag cat ccc aag aac ttt ggc ctg atc gca 1518
147 Phe Arg Glu Lys Phe Met Gln His Pro Lys Asn Phe Gly Leu Ile Ala
148 440 445 450
150 tca ttc ctg gag agg aag aca gtg gct gag tgc gtc ctc tat tac tac 1566
151 Ser Phe Leu Glu Arg Lys Thr Val Ala Glu Cys Val Leu Tyr Tyr Tyr
152 455 460 465 470
154 ctg act aag aag aat gag aac tat aag agc ctg gtg aga cgg agc tat 1614
155 Leu Thr Lys Lys Asn Glu Asn Tyr Lys Ser Leu Val Arg Arg Ser Tyr
156 475 480 485
158 cgg cgc cgc ggc aag agc cag cag caa caa cag cag cag cag cag cag 1662
159 Arg Arg Arg Gly Lys Ser Gln Gln Gln Gln Gln Gln Gln Gln Gln
160 490 495 500
162 cag cag cag cag cag cag cag ccc atg ccc cgc agc agc cag gag gag 1710
163 Gln Gln Gln Gln Gln Gln Gln Pro Met Pro Arg Ser Ser Gln Glu Glu
164 505 510 515
166 aaa gat gag aag gag aag gaa aag gag gcg gag aag gag gag gag aag 1758
167 Lys Asp Glu Lys Glu Lys Glu Lys Glu Ala Glu Lys Glu Glu Glu Lys
168 520 525 530
170 ccg gag gtg gag aac gac aag gaa gac ctc ctc aag gag aag aca gac 1806
171 Pro Glu Val Glu Asn Asp Lys Glu Asp Leu Leu Lys Glu Lys Thr Asp
172 535 540 545 550
174 gac acc tca ggg gag gac aac gac gag aag gag gct gtg gcc tcc aaa 1854
175 Asp Thr Ser Gly Glu Asp Asn Asp Glu Lys Glu Ala Val Ala Ser Lys
176 555 560 565
178 ggc cgc aaa act gcc aac agc cag gga aga cgc aaa ggc cgc atc acc 1902
179 Gly Arg Lys Thr Ala Asn Ser Gln Gly Arg Arg Lys Gly Arg Ile Thr
180 570 575 580
182 cgc tca atg gct aat gag gcc aac agc gag gag gcc atc acc ccc cag 1950
183 Arg Ser Met Ala Asn Glu Ala Asn Ser Glu Glu Ala Ile Thr Pro Gln
184 585 590 595
186 cag agc gcc gag ctg gcc tcc atg gag ctg aat gag agt tct cgc tgg 1998
187 Gln Ser Ala Glu Leu Ala Ser Met Glu Leu Asn Glu Ser Ser Arg Trp
188 600 605 610
190 aca gaa gaa gaa atg gaa aca gcc aag aaa ggt ctc ctg gaa cac ggc 2046
191 Thr Glu Glu Glu Met Glu Thr Ala Lys Lys Gly Leu Leu Glu His Gly
192 615 620 625 630
194 cgc aac tgg tcg gcc atc gcc cgg atg gtg ggc tcc aag act gtg tcg 2094

```

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/819,104A

TIME: 15:03:47

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04122002\I819104A.raw

195	Arg	Asn	Trp	Ser	Ala	Ile	Ala	Arg	Met	Val	Gly	Ser	Lys	Thr	Val	Ser	
196					635					640					645		
198	cag	tgt	aag	aac	ttc	tac	ttc	aac	tac	aag	aag	agg	cag	aac	ctc	gat	2142
199	Gln	Cys	Lys	Asn	Phe	Tyr	Phe	Asn	Tyr	Lys	Lys	Arg	Gln	Asn	Leu	Asp	
200				650				655					660				
202	gag	atc	ttg	cag	cag	cac	aag	ctg	aag	atg	gag	aag	gag	agg	aac	gcg	2190
203	Glu	Ile	Leu	Gln	Gln	His	Lys	Leu	Lys	Met	Glu	Lys	Glu	Arg	Asn	Ala	
204				665				670					675				
206	cgg	agg	aag	aag	aag	aaa	gcg	ccg	gcg	gcg	gcc	agc	gag	gag	gct	gca	2238
207	Arg	Arg	Lys	Lys	Lys	Lys	Ala	Pro	Ala	Ala	Ala	Ser	Glu	Glu	Ala	Ala	
208		680					685					690					
210	ttc	ccg	ccc	gtg	gtg	gag	gat	gag	gag	atg	gag	gcg	tcg	ggc	gtg	acg	2286
211	Phe	Pro	Pro	Val	Val	Glu	Asp	Glu	Glu	Met	Glu	Ala	Ser	Gly	Val	Thr	
212	695					700				705					710		
214	gga	aat	gag	gag	gag	atg	gtg	gag	gag	gct	gaa	gcc	act	gtc	aac	aac	2334
215	Gly	Asn	Glu	Glu	Glu	Met	Val	Glu	Glu	Ala	Glu	Ala	Thr	Val	Asn	Asn	
216					715					720					725		
218	agc	tca	gac	acc	gag	agc	atc	ccc	tct	cct	cac	act	gag	gcc	gcc	aag	2382
219	Ser	Ser	Asp	Thr	Glu	Ser	Ile	Pro	Ser	Pro	His	Thr	Glu	Ala	Ala	Lys	
220				730				735					740				
222	gac	aca	ggg	cag	aat	ggg	ccc	aag	ccc	cca	gcc	acc	ctg	ggc	gcc	gac	2430
223	Asp	Thr	Gly	Gln	Asn	Gly	Pro	Lys	Pro	Pro	Ala	Thr	Leu	Gly	Ala	Asp	
224		745					750					755					
226	ggg	cca	ccc	cca	ggg	cca	ccc	acc	cca	cca	ccg	gag	gac	atc	ccg	gcc	2478
227	Gly	Pro	Pro	Pro	Gly	Pro	Pro	Thr	Pro	Pro	Pro	Glu	Asp	Ile	Pro	Ala	
228		760					765					770					
230	ccc	act	gag	tcc	acc	ccg	gcc	tct	gaa	gcc	acc	tta	gcc	cct	acg	ccc	2526
231	Pro	Thr	Glu	Ser	Thr	Pro	Ala	Ser	Glu	Ala	Thr	Leu	Ala	Pro	Thr	Pro	
232	775					780				785					790		
234	cca	cca	gca	ccc	cca	ttt	ccc	tct	tca	cct	cct	cct	gtg	gtc	ccc	aag	2574
235	Pro	Pro	Ala	Pro	Pro	Phe	Pro	Ser	Ser	Pro	Pro	Pro	Val	Val	Pro	Lys	
236				795				800					805				
238	gag	gag	aag	gag	gag	gag	acc	gca	gca	gcg	ccc	cca	gtg	gag	gag	ggg	2622
239	Glu	Glu	Lys	Glu	Glu	Glu	Thr	Ala	Ala	Ala	Pro	Pro	Val	Glu	Glu	Gly	
240			810					815					820				
242	gag	gag	cag	aag	ccc	ccc	gcg	gct	gag	gag	ctg	gca	gtg	gac	aca	ggg	2670
243	Glu	Glu	Gln	Lys	Pro	Pro	Ala	Ala	Glu	Glu	Leu	Ala	Val	Asp	Thr	Gly	
244			825				830						835				
246	aag	gcc	gag	gag	ccc	gtc	aag	agc	gag	tgc	acg	gag	gaa	gcc	gag	gag	2718
247	Lys	Ala	Glu	Glu	Pro	Val	Lys	Ser	Glu	Cys	Thr	Glu	Glu	Ala	Glu	Glu	
248		840					845					850					
250	ggg	ccg	gcc	aag	ggc	aag	gac	gcg	gag	gcc	gct	gag	gcc	acg	gcc	gag	2766
251	Gly	Pro	Ala	Lys	Gly	Lys	Asp	Ala	Glu	Ala	Ala	Glu	Ala	Thr	Ala	Glu	
252	855				860					865					870		
254	agg	gcg	ctc	aag	gca	gag	aag	aag	gag	ggc	ggg	agc	ggc	agg	gcc	acc	2814
255	Arg	Ala	Leu	Lys	Ala	Glu	Lys	Lys	Glu	Gly	Gly	Ser	Gly	Arg	Ala	Thr	
256				875				880					885				
258	aca	gcc	aag	agc	tcg	ggc	gcc	ccc	cag	gac	agc	gac	tcc	agt	gcc	acc	2862
259	Thr	Ala	Lys	Ser	Ser	Gly	Ala	Pro	Gln	Asp	Ser	Asp	Ser	Ser	Ala	Thr	

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/819,104A

TIME: 15:03:47

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04122002\I819104A.raw

260		890		895		900		
262	tgc agt gca	gag gtg gat	gag gcc gag	ggc ggc gac	aag aac cgg	2910		
263	Cys Ser Ala	Asp Glu Val	Asp Glu Ala	Glu Gly Gly	Asp Lys Asn	Arg		
264		905		910		915		
266	ctg ctg tcc	cca agg ccc	agc ctc ctc	acc ccg act	ggc gac ccc	cgg	2958	
267	Leu Leu Ser	Pro Arg Pro	Ser Leu Leu	Thr Pro Thr	Gly Asp Pro	Arg		
268		920		925		930		
270	gcc aat gcc	tca ccc cag	aag cca ctg	gac ctg aag	cag ctg aag	cag	3006	
271	Ala Asn Ala	Ser Pro Gln	Lys Pro Leu	Asp Leu Lys	Gln Leu Lys	Gln		
272	935		940		945	950		
274	cga gcg gct	gcc atc ccc	ccc atc cag	gtc acc aaa	gtc cat gag	ccc	3054	
275	Arg Ala Ala	Ile Pro Pro	Ile Gln Val	Thr Lys Val	His Glu Pro			
276		955		960		965		
278	ccc cgg gag	gac gca gct	ccc acc aag	cca gct ccc	cca gcc cca	ccg	3102	
279	Pro Arg Glu	Asp Ala Ala	Pro Thr Lys	Pro Ala Pro	Pro Ala Pro	Pro		
280		970		975		980		
282	cca ccg caa	aac ctg cag	ccg gag agc	gac gcc cct	cag cag cct	ggc	3150	
283	Pro Pro Gln	Asn Leu Gln	Pro Glu Ser	Asp Ala Pro	Gln Gln Pro	Gly		
284		985		990		995		
286	agc agc ccc	cgg ggc aag	agc agg agc	ccg gca ccc	ccc gcc gac	aag	3198	
287	Ser Ser Pro	Arg Gly Lys	Ser Arg Ser	Pro Ala Pro	Pro Ala Asp	Lys		
288		1000		1005		1010		
290	gag gca gag	aag cct gtg	ttc ttc cca	gcc ttc gca	gcc gag gcc	cag	3246	
291	Glu Ala Glu	Lys Pro Val	Phe Phe Pro	Ala Phe Ala	Ala Glu Ala	Gln		
292	1015		1020		1025	1030		
294	aag ctg cct	ggg gac ccc	cct tgc tgg	act tcc ggc	ctg ccc ttc	ccc	3294	
295	Lys Leu Pro	Gly Asp Pro	Pro Cys Trp	Thr Ser Gly	Leu Pro Phe	Pro		
296		1035		1040		1045		
298	gtg ccc ccc	cgt gag gtg	atc aag gcc	tcc ccg cat	gcc ccg gac	ccc	3342	
299	Val Pro Pro	Arg Glu Val	Ile Lys Ala	Ser Pro His	Ala Pro Asp	Pro		
300		1050		1055		1060		
302	tca gcc ttc	tcc tac gct	cca cct ggt	cac cca ctg	ccc ctg ggc	ctc	3390	
303	Ser Ala Phe	Ser Tyr Ala	Pro Pro Gly	His Pro Leu	Pro Leu Gly	Leu		
304		1065		1070		1075		
306	cat gac act	gcc cgg ccc	gtc ctg ccg	cgc cca ccc	acc atc tcc	aac	3438	
307	His Asp Thr	Ala Arg Pro	Val Leu Pro	Arg Pro Thr	Ile Ser Asn			
308		1080		1085		1090		
310	ccg cct ccc	ctc atc tcc	tct gcc aag	cac ccc agc	gtc ctc gag	agg	3486	
311	Pro Pro Pro	Leu Ile Ser	Ser Ala Lys	His Pro Ser	Val Leu Glu	Arg		
312	1095		1100		1105	1110		
314	caa ata ggt	gcc atc tcc	caa gga atg	tgc gtc cag	ctc cac gtc	ccg	3534	
315	Gln Ile Gly	Ala Ile Ser	Gln Gly Met	Ser Val Gln	Leu His Val	Pro		
316		1115		1120		1125		
318	tac tca gag	cat gcc aag	gcc ccg gtg	ggc cct gtc	acc atg ggg	ctg	3582	
319	Tyr Ser Glu	His Ala Lys	Ala Pro Val	Gly Pro Val	Thr Met Gly	Leu		
320		1130		1135		1140		
322	ccc ctg ccc	atg gac ccc	aaa aag ctg	gca ccc ttc	agc gga gtg	aag	3630	
323	Pro Leu Pro	Met Asp Pro	Lys Lys Leu	Ala Pro Phe	Ser Gly Val	Lys		
324		1145		1150		1155		

VERIFICATION SUMMARY

DATE: 04/12/2002

PATENT APPLICATION: US/09/819,104A

TIME: 15:03:49

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\04122002\I819104A.raw